Real-time imaging of co-transcriptional splicing reveals a kinetic model that reduces noise: implications for alternative splicing regulation.

Published May 30, 2011

Real-time imaging of cotranscriptional splicing reveals a kinetic model that reduces noise: implications for alternative splicing regulation

Ute Schmidt,[1] Eugenia Basyuk,[1] Marie-Cécile Robert,[1] Minoru Yoshida,[2] Jean-Philippe Villemin,[3] Didier Auboeuf,[3] Stuart Aitken,[4,5] and Edouard Bertrand[1]

[1]Institut de Génétique Moléculaire de Montpellier, Centre National de la Recherche Scientifique UMR 5535, 34293 Montpellier Cedex 5, France
[2]Chemical Genomics Research Group and Chemical Genetics Laboratory, Institute of Physical and Chemical Research Institute, Saitama 351-0198, Japan
[3]Centre de Recherche en Cancérologie de Lyon, UMR Institut National de la Santé et de la Recherche Médicale U1052, Centre National de la Recherche Scientifique U5286, Centre Léon Bérard, 69008 Lyon, France
[4]School of Informatics and [5]Centre for Systems Biology at Edinburgh, University of Edinburgh, Edinburgh, Scotland EH9 3JD, UK

Splicing is a key process that expands the coding capacity of genomes. Its kinetics remain poorly characterized, and the distribution of splicing time caused by the stochasticity of single splicing events is expected to affect regulation efficiency. We conducted a small-scale survey on 40 introns in human cells and observed that most were spliced cotranscriptionally. Consequently, we constructed a reporter system that splices cotranscriptionally and can be monitored in live cells and in real time through the use of MS2–GFP. All small nuclear ribonucleoproteins (snRNPs) are loaded on nascent pre-mRNAs, and spliceostatin A inhibits splicing but not snRNP recruitment. Intron removal occurs in minutes and is best described by a model where several successive steps are rate limiting. Each pre-mRNA molecule is predicted to require a similar time to splice, reducing kinetic noise and improving the regulation of alternative splicing. This model is relevant to other kinetically controlled processes acting on few molecules.

Introduction

Alternative splicing is a process that alters the coding capacity of genomes (Sultan et al., 2008; Wang et al., 2008), and it thus occupies a unique place in the control of gene expression (for review see Blencowe, 2006). Splicing is a two-step chemical reaction performed by the spliceosome, a complex RNP machine that assembles anew on every intron (for review see Wahl et al., 2009). It is composed of five RNAs and about a hundred proteins, and in vitro studies have defined an ordered assembly pathway. First, U1 small nuclear RNA (snRNA) binds the 5′ splice site, whereas the branch point, poly-pyrimidine tract, and 3′ splice site are recognized by SF1, U2AF65, and U2AF35, respectively. U2 snRNP is then loaded in an ATP-dependent step to form the A complex. The U4/U5/U6 tri-snRNP further joins to form the B complex, which loses U1 and U4 and undergoes remodeling to become competent for catalysis. After the first trans-esterification reaction, the composition of the spliceosome is again modified, yielding the C complex that performs the second trans-esterification. Finally, post-spliceosomes are disassembled, their components recycled and introns are degraded.

In vivo, splicing is intimately connected with transcription (for review see Pandit et al., 2008). Studies on a few mRNAs have shown that splicing can occur cotranscriptionally (Beyer and Osheim, 1988; Baurén and Wieslander, 1994; Kiseleva et al., 1994; Zhang et al., 1994; Pandya-Jones and Black 2009; de la Mata et al., 2010). However, recent studies in yeast have indicated that the relative kinetics of splicing versus transcription determine whether an intron splices cotranscriptionally or not (Tardiff et al., 2006; Carrillo-Oesterreich et al., 2010). We currently lack a genome-wide view of co- and posttranscriptional splicing in higher eukaryotes. However, there are many links between splicing and transcription. First, the RNA polymerase II recruits splicing factors to promote cotranscriptional

Correspondence to Edouard Bertrand: Edouard.Bertrand@igmm.cnrs.fr; or Eugenia Basyuk: Eugenia.Basyuk@igmm.cnrs.fr

Abbreviations used in this paper: AIC, Akaike information criterion; snRNA, small nuclear RNA; SSA, spliceostatin A.

Supplemental Material can be found at:
http://jcb.rupress.org/content/suppl/2011/05/26/jcb.201009012.DC1.html
http://jcb.rupress.org/content/suppl/2011/05/27/jcb.201009012.DC2.html

The Rockefeller University Press $30.00
J. Cell Biol. Vol. 193 No. 5 819–829
www.jcb.org/cgi/doi/10.1083/jcb.201009012

Published May 30, 2011

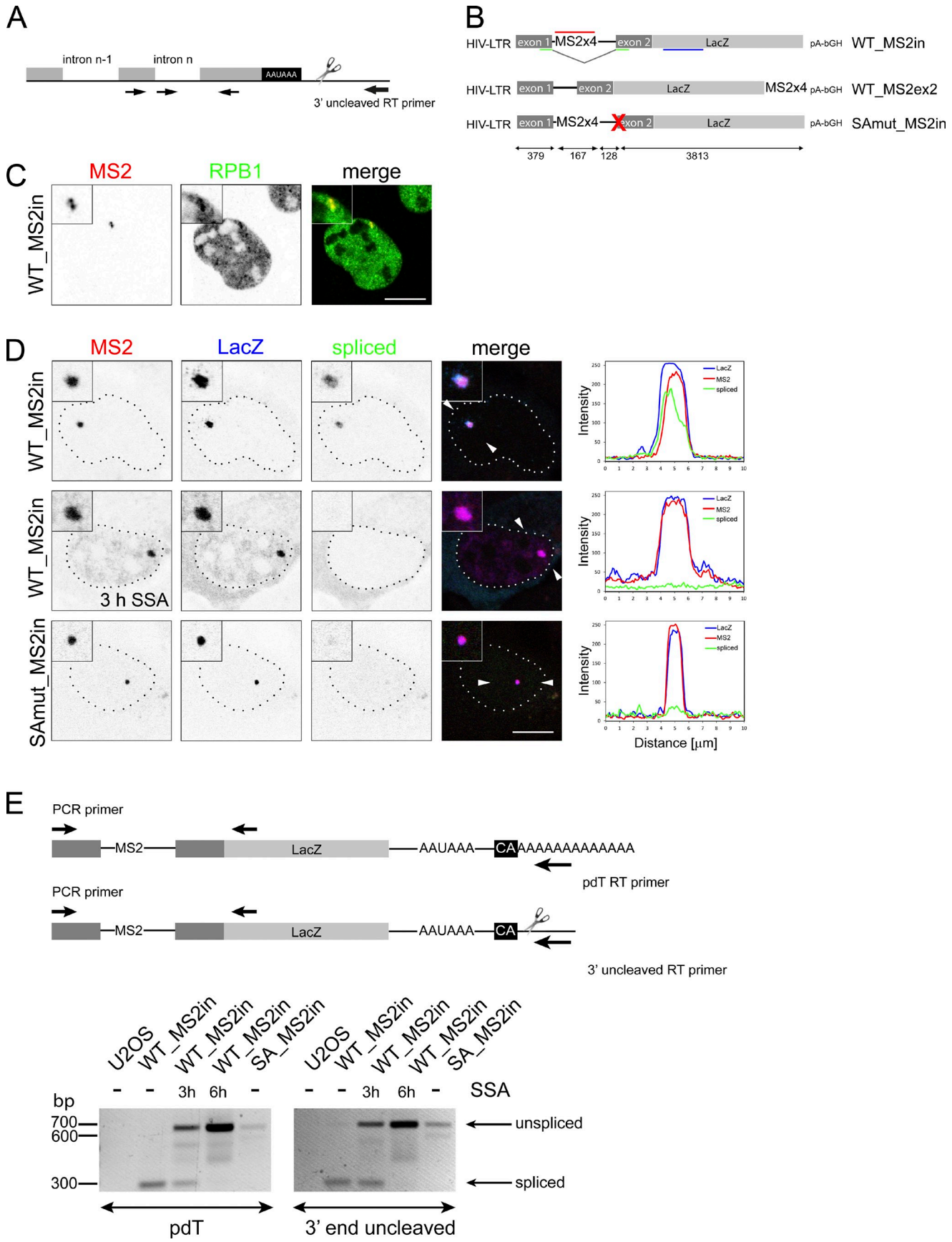

Figure 1. **Co-transcriptional splicing and characterization of the MINX reporter cell lines.** (A) RT-PCR assay to determine whether introns splice cotranscriptionally or not. A hypothetical gene is schematized with the position of the primers used for RT and the competitive PCR. (B) Scheme of the MINX constructs. SAmut_MS2in contains a single AG to GG mutation at the splice acceptor site. The green, red, and blue bars represent the positions of the FISH probes,

Published May 30, 2011

splicing (Kim et al., 1997; McCracken et al., 1997). Second, transcription elongation rates have been shown to regulate alternative splicing (Roberts et al., 1998). Indeed, pause sites or slowing down RNA polymerase II delay the synthesis of competing splice sites and provide a kinetic way to regulate alternative splicing (de la Mata et al., 2003). For instance, when cells are exposed to UV light, RNA polymerase II becomes hyperphosphorylated, and its elongation rate decreases to induce global changes in alternative splicing (Muñoz et al., 2009). Understanding this regulation requires the kinetic characterization of splicing and transcription. In particular, the time taken to splice a single pre-mRNA should match the elongation rate of the polymerase that transcribes it, and the stochasticity of single splicing events is expected to reduce regulation efficiency. However, splicing kinetics have not been characterized in detail, mainly because current methods do not provide a high temporal resolution.

Here, we have created a reporter gene that splices cotranscriptionally and can be monitored in real time and in live cells. We provide a detailed kinetic analysis of splicing in vivo, and propose a model that explains how the variability in splicing time, which we refer to as kinetic noise, is reduced to override the stochasticity of single splicing events, thereby improving splicing regulation.

Results and discussion

Cotranscriptional splicing is frequent in human cells

Although cotranscriptional splicing has been demonstrated in human cells, it is not clear whether this is a predominant mode of splicing or not. To address this question, we chose a set of 40 constitutive and alternatively spliced human introns and determined with a quantitative assay whether they splice cotranscriptionally or not. To detect nascent RNAs still attached to chromatin, we used an RT primer situated beyond the 3′ cleavage and polyadenylation site. Then, we performed competitive RT-PCR assays with primers that amplified simultaneously the spliced and the unspliced RNAs for the last two introns of the selected genes (Fig. 1 A). Because the distance the reverse transcription can extend is limited, we selected introns that were <3 kb from the cleavage and polyadenylation site. Remarkably, most introns were spliced completely or nearly completely before 3′ end maturation (Table I). This was also the case for four large introns (>1 kb) and for introns that showed alternative splicing. Only one intron was found to splice predominantly posttranscriptionally. Although this set of introns is small and biased for short introns, these data nevertheless suggest that cotranscriptional splicing is frequent in human cells. We thus developed a model system that splices cotranscriptionally and that can be monitored in living cells.

A reporter system to analyze splicing in live cells and in real time

We and others have previously analyzed mRNA biogenesis in real time, using MS2-tagged constructs integrated in the genome of mammalian cells and an MS2–GFP fusion protein (Fusco et al., 2003; Boireau et al., 2007; Darzacq et al., 2007). In this setting, the transcription site of the reporter gene is visible as a bright spot in the nucleoplasm and it is possible to measure the kinetics of transcription and 3′ end processing by photobleaching the MS2–GFP protein bound to nascent mRNAs (Boireau et al., 2007; Darzacq et al., 2007). Here, we introduced the MS2 binding sites within the MINX intron to monitor the splicing reaction (reporter WT_MS2in, Fig. 1 B). This small intron derives from the adenovirus genome and possesses strong splicing signals (Zillmann et al., 1988). We used only four MS2 binding sites instead of 24 to minimize possible perturbations. To increase the chances that splicing occurs cotranscriptionally, we placed the tagged intron upstream of the LacZ gene (3 kb long), and we used the cleavage and polyadenylation sequence of the bovine growth hormone gene (*bGH*), which is processed at a slow rate (0.2 min^{-1}, as opposed to 1 min^{-1} for the HIV-1 polyA signals; Boireau et al., 2007). Transcription was driven by the HIV-1 LTR and was induced by transfecting a Tat expression vector.

To compare the residency time of the intron with the one of the nascent mRNAs, we derived two other constructs (Fig. 1 B). In the first, the MS2 sites were inserted in exon 2 (WT_MS2ex2). In the second, we inhibited splicing with a point mutation in the splice acceptor sequence (AG→GG, SAmut_MS2in).

The MS2-tagged MINX intron splices cotranscriptionally

The various reporters were transfected into U2OS cells, and stable clones were selected and characterized. These clones contained 18–20 copies of the reporter gene, and we first used fluorescent oligonucleotide probes to test whether splicing of the MS2-tagged MINX intron occurs at its transcription site. Three probe sets were used (Fig. 1 B). One hybridized to the intronic MS2 sites, one was specific for the spliced mRNA, and one hybridized at the beginning of the LacZ gene in exon 2. The MS2 probes labeled a bright spot in the nucleoplasm. This spot colocalized with a local accumulation of the large subunit of RNA polymerase II (RPB1) and thus corresponded to the

and the distance is indicated in nucleotides. (C) RNA polymerase II accumulates at the reporter transcription site. WT_MS2in stable cells were transfected with Tat and processed for in situ hybridization with an MS2 probe (red), and for immunofluorescence against the large subunit of RNA polymerase II (RPB1, green). Insets show enlarged images. Bar, 10 µm. (D) WT_MS2in is spliced cotranscriptionally. WT_MS2in or SAmut_MS2in cells were treated with SSA when indicated, and processed for in situ hybridization with probes recognizing the intron (MS2, red), exon 2 (LacZ, blue), or the spliced junction (spliced, green). Bar, 10 µm. Right panels show intensity line scans of the images, using the lines defined by the arrowheads. (E) Quantification of cotranscriptional splicing by RT-PCR. The indicated cell lines were treated as in D, and total RNAs were extracted and RT-PCR amplified with the indicated primers (see scheme, the scissors depict the 3′ end cleavage site). The position of the primers used for reverse transcription are also indicated. pdT, oligo dT primer; 3′ uncleaved primer, primer located downstream the polyadenylation site. Spliced RNAs: 317 bp; unspliced RNAs: 671 bp.

Published May 30, 2011

Table I. **Small-scale study of co- and posttranscriptional splicing patterns**

Gene	Co-transcriptional splicing, intron n − 1	Size intron	Type	Cotranscriptional splicing, intron n (last)	Size intron	Type
		bp			*bp*	
Constitutive introns						
ILK	>90%	112	NA	>85%	183	NA
RPL9	>95%	214	NA	>80%	307	NA
PSMB6	100%	197	NA	100%	126	NA
UTP15	100%	166	NA	>90%	501	NA
ATXN7L3	100%	93	NA	>90%	105	NA
EDC4	>85%	390	NA	>90%	224	NA
PSENEN	100%	427	NA	>75%	184	NA
ALKBH7	ND	88	NA	100%	221	NA
ROGDI	100%	97	NA	>85%	212	NA
G6PC3	>90%	222	NA	ND	NA	NA
LASS2	100%	126	NA	ND	NA	NA
TK1	>95%	100	NA	ND	NA	NA
EWSR	ND	NA	NA	>80%	3,600	NA
Alternatively spliced introns						
ARL6IP	100%	395	Constitutive	>90%	82	Intron inclusion (<5%)
SYNGR2	100%	456	Constitutive	100%	90	Intron inclusion (<5%)
FLAD1	85%	2,185	Constitutive	>90%	116	Intron inclusion (5%)
ARRDC1	80%	84	Constitutive	70%	74	Intron inclusion (30%)
FAM132B	80%	1,272	Constitutive	>95%	66	Intron inclusion (60%)
TUBB2C	100%	82	Constitutive	100%	475	Alternative 5′ and 3′ ss (>95%)
CLDN7	ND	NA	NA	<10%	285	Exon cassette (90%)
TUBA1A	ND	NA	NA	>90%	621	Exon cassette (>95%)
LMNA	>80%	1,256	Exon cassette (>95%)	100%	323	Constitutive

Frequences of alternative splicing events in U20S cells are shown in parentheses. NA, not applicable.

mRNA transcription site (Fig. 1 C). When the three probes were used simultaneously, the spliced mRNA showed a signal similar to that of LacZ, labeling both the cytoplasm and the transcription site, thus suggesting that splicing is cotranscriptional (Fig. 1 D and not depicted). To confirm this data, we treated cells with spliceostatin A (SSA), a small molecule that inhibits splicing (Kaida et al., 2007). 3 h after addition of the drug, the spliced probe failed to detect the transcription site, whereas it still labeled the cytoplasmic mRNAs that were synthesized before addition of the drug (Fig. 1 D). Conversely, the signal for the intronic MS2 probe that was restricted to the transcription site became present throughout the nucleoplasm in a speckled pattern, and was even detected in the cytoplasm in a small fraction of the cells.

The results obtained with SSA were confirmed with the mutant reporter gene that lacked a functional 3′ splice site (Fig. 1 D): the exonic LacZ probes labeled a bright spot in the nucleoplasm and showed a diffuse staining in the nucleoplasm and the cytoplasm; the intronic MS2 probes labeled the transcription site, the nucleoplasm, and the cytoplasm; and the spliced probe did not detect any signal. Altogether, these results demonstrated that a fraction of the wild-type pre-mRNAs was spliced at their transcription site.

We then used an RT-PCR assay to determine the amount of pre-mRNA that spliced co- and posttranscriptionally. We primed the reverse transcription with oligos specific for uncleaved, nascent pre-mRNAs or an oligo-dT primer to detect released, polyadenylated pre-mRNAs (Fig. 1 E). The wild-type reporter showed a single PCR product that corresponded to the spliced RNAs. This occurred for both nascent and polyadenylated mRNAs, demonstrating that splicing of the reporter is completed before it is cleaved and polyadenylated (Fig. 1 E). 3 h after SSA addition, polyadenylated and nascent RNAs showed both spliced and unspliced species, and the spliced band disappeared after 24 h of treatment. The splice acceptor mutant reporter showed predominantly unspliced pre-mRNAs, with a minor fraction that was spliced at a cryptic downstream site (unpublished data). Altogether these experiments firmly established that the wild-type reporter was spliced cotranscriptionally, and that this was blocked by the addition of SSA or by mutating the 3′ splice site.

Recruitment of splicing factors to the MINX transcription site

Next, we determined whether splicing factors were recruited to the MINX transcription site using two-color FISH. In untreated cells, all snRNAs were detected at the transcription site, but accumulation was weak (Fig. 2 A). In contrast, in cells treated with SSA for 3 h, all the snRNA accumulated at the MINX transcription site at high levels (Fig. 2 B).

Published May 30, 2011

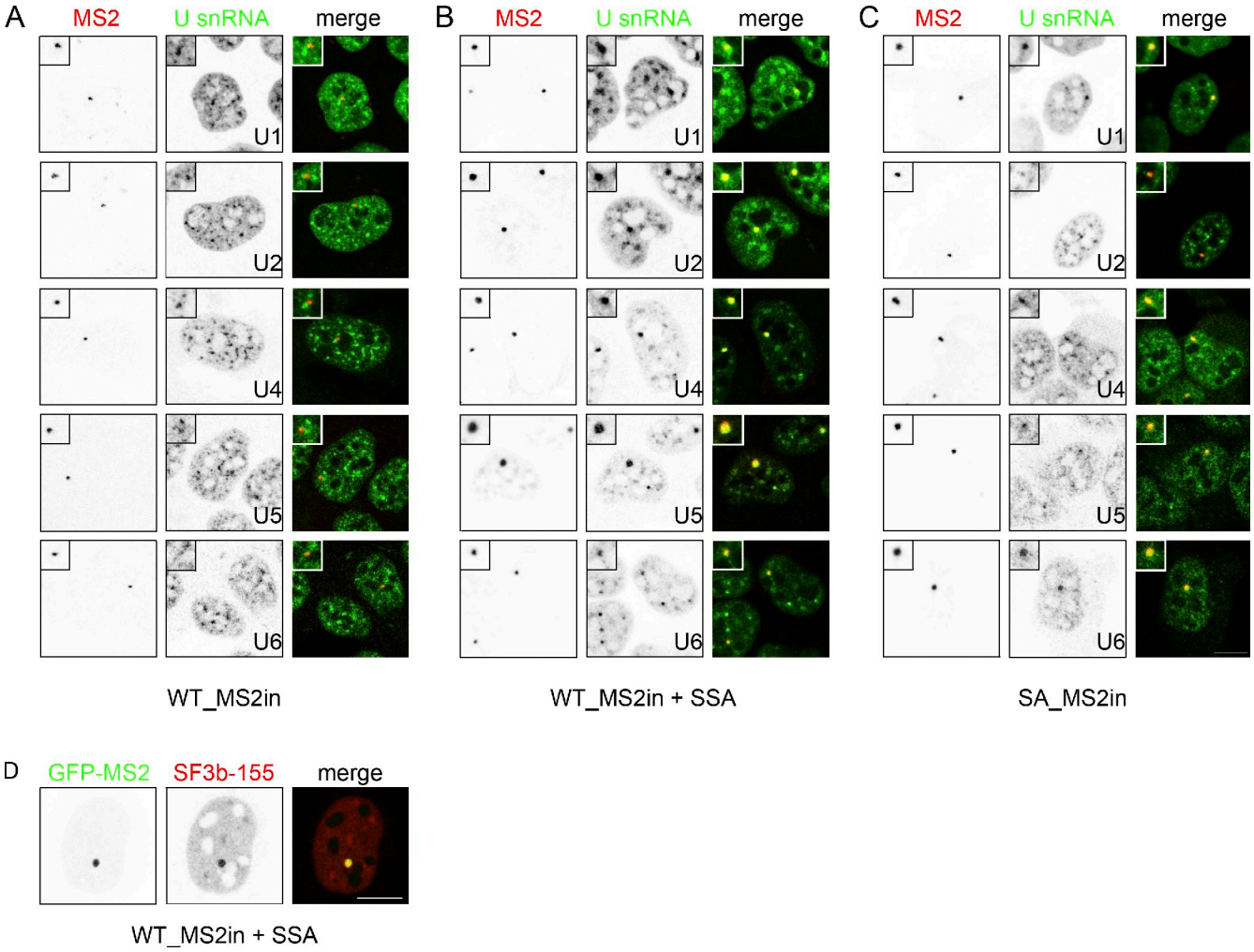

Figure 2. **Recruitment of spliceosomal components to the MINX transcription site.** (A–C) Recruitment of snRNPs. WT_MS2in and SAmut_MS2in cells were transfected with Tat vector and processed for multicolor FISH experiments with probes against LacZ in exon 2 (red), and each of the snRNAs (green). In B, cells were treated with SSA. Insets show enlarged views. (D) SF3b155 recruitment in SSA-treated cells. WT_MS2in cells were transfected with vectors expressing Tat and MS2–GFP (green), treated with SSA for 3 h, and processed for immunofluorescence against SF3B155 (red). Bars, 10 µm.

This suggested that the weak accumulation observed in untreated cells was caused by the high efficiency of the splicing reaction, as the snRNPs resided only shortly on the pre-mRNA to complete splicing. This is consistent with previous studies of spliceosome assembly in fixed and live cells, which estimated that spliceosomal components resided on pre-mRNAs for 15–30 s on average (Wetterberg et al., 2001; Huranová et al., 2010). Interestingly, SF3b155, the known spliceostatin target, was also recruited in cells treated with the drug (Fig. 2 D). Our data thus suggest that SSA does not block spliceosome assembly, but inhibits spliceosome activation, which is in agreement with in vitro data (Roybal and Jurica, 2010). We next analyzed the splice acceptor mutant. Similar to SSA-treated cells, we found that all snRNAs were enriched at the transcription site of the mutant reporter, including U1 and U4 (Fig. 2 C). Altogether, this demonstrated that the spliceosome was assembled cotranscriptionally on the MINX reporter, and that SSA and mutating the splice acceptor site impaired splicing catalysis but not recruitment of spliceosomal components.

Splicing kinetics and modeling of the reaction

The presence of the MS2 tag within the intron allows its visualization in live cells (Fusco et al., 2003). To determine the half-life of the intron, MS2–GFP was bleached at the MINX transcription site (Fig. 3). As the measurements improve when the signal of the transcription sites increases over the nucleoplasmic background, a nuclear export signal was added to the original nuclear MS2–GFPnls protein. This way, the MS2–GFPnls-nes displays lower nucleoplasmic signals and allows for a better visualization of nuclear RNAs (Fig. 3). Using this technique, we measured a half-life of 105 s for the MS2–GFP labeled intron, significantly shorter than the 165 s for exon 2.

It was previously demonstrated that the binding of MS2–GFP to the MS2 stem loops is relatively stable in live cells, as little dissociation occurs in the first 10 min of a FRAP experiment. In contrast, diffusion of free MS2–GFP is rapid (Boireau et al., 2007). Indeed, recovery of the free pool of MS2–GFP was complete 10 s after the bleach, at which time there was no visible recovery of the transcription sites (Fig. 3 A). Diffusion and

Published May 30, 2011

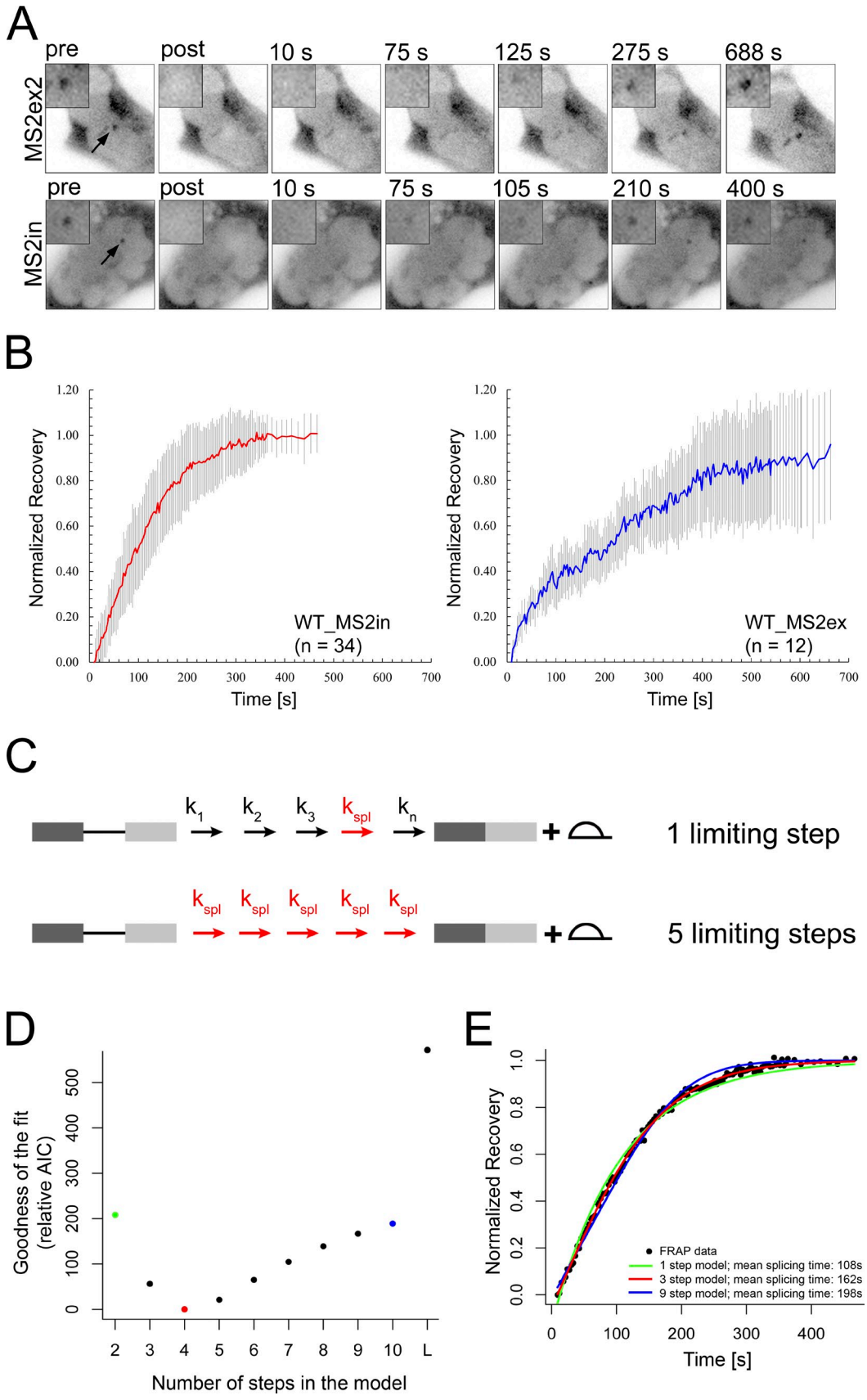

Figure 3. **Analysis of splicing kinetics by FRAP.** (A) Time series of FRAP experiments. WT_MS2ex2 (top) or WT_MS2in (bottom) cells were transfected with vectors expressing Tat and MS2–GFPnsl_nes. Transcription sites were bleached and 3D images were captured for 9–11 min. The time after the bleach is indicated in seconds. pre, prebleach image; post, post-bleach image. The arrow points to the transcription site. Bar, 10 µm. (B) FRAP recovery plots. Normalized fluorescence intensity at the transcription site is plotted as a function of time (±SEM). Time is in seconds, and the zero time point corresponds

Published May 30, 2011

dissociation of MS2–GFP were thus neglected in the interpretation of the FRAP curves. Intron turnover included transcription from the MS2 repeat to the end of the MINX intron (230 nucleotides), splicing, and diffusion and/or degradation of the free intron. Given the estimated rates of RNA polymerase II (1.8–4 kb/min; Boireau et al., 2007; Darzacq et al., 2007; Wada et al., 2009), transcription to the intron end should take between 3 and 7.7 s. Furthermore, Northern blots probed with MS2 repeats detected only full-length or nearly full-length mRNAs, rather than the free intron (Fig. S1), which indicates that intron degradation is rapid and not rate limiting.

Spliceosomes must assemble de novo on each intron. Given the number of components that bind and become assembled into a functional enzyme, this process includes a large number of elementary steps. Modeling this complicated series of events is currently out of reach. Instead, we decided to determine the number of steps that are rate-limiting for the reaction by fitting the FRAP curve with models containing a single, two, three, and up to nine consecutive limiting steps (Fig. 3 C). If a single step in the assembly or catalysis pathway is rate limiting, it should impose its kinetic on the whole process, and the FRAP curve should adopt a simple exponential shape. In contrast, if there are several successive rate-limiting steps, different kinetics will be observed (Fig. 3 E), and this is a key point for alternative splicing regulation (see the following paragraph).

We generated a full model that describes transcription and splicing, but to fit the data, we derived a simplified model where the time for the polymerase to transcribe to the 3′ splice site is represented by a time offset (β), which was constrained to 3–7.7 s (Fig. S2). We found a general analytical solution for the simplified model, and models with 1 and up to 9 limiting steps were thus fitted to the FRAP data by finding the optimal values of the step rate (α) and the time offset (β; see Materials and methods and Fig. S2). Notably, all models have an identical number of variables (the step rate, time offset, and a scaling constant), allowing for a fair comparison between them. A linear model was also included as a null hypothesis (Fig. S2). The Akaike information criterion (AIC) values for the optimal choice of α show that the fits improve from one to two limiting steps, is optimal for a three step model, and then worsens as the number of steps further increases (Fig. 3, D and E). Other indicators of the goodness of the fits, such as the distribution of the residuals and the reduced χ^2 also indicate that the three-step model is the best and that models with one step or more than five steps are unlikely (Table S1). In addition, the optimal parameters found for the three-step model allowed for an accurate description of the FRAP curve using the full model (see Fig. S2 for details).

A three-step model indicated a mean splicing time of 162 s for MINX. Splicing rates have been analyzed by various methods and in several organisms: by EM in *Chironomus tentans* and *Drosophila melanogaster* (Beyer and Osheim, 1988; Kiseleva et al., 1994; Wetterberg et al., 2001), by transcriptional pulse-chase in mammals (Audibert et al., 2002; Singh and Padgett, 2009), and by chromatin immunoprecipitation (ChIP) and other assays in yeast (Tardiff et al., 2006). Although these methods lack the temporal resolution provided by the FRAP assay, the estimated splicing rates are in good agreement with our MINX data. Indeed, splicing was estimated to occur with a half-time of 0.4–7 min in mammalian cells and to be complete in 3 min in *Drosophila*, in 0.5–5 min in *C. tentans*, and in <1 min in yeast.

Kinetics and noise of splicing

Next, we used values of the best fits to simulate the production of the spliced product (Figs. 4 and S2 G). With a single limiting step, the appearance of the spliced mRNA follows an exponential curve. However, with three limiting steps, the appearance of the spliced message has a more sigmoidal shape (Fig. 4 A). These kinetics can be understood from the viewpoint of a stochastic model. In the case of a single limiting step, some molecules will be spliced quickly, and others slowly, because of the stochasticity of the single step. In contrast, if there are several successive rate-limiting steps occurring at a similar rate, the stochastic nature of each event will be averaged by the larger number of steps involved. Each molecule will take a similar time to splice, thereby mimicking a deterministic process. Thus, the number of steps influences the distribution of splicing times around the mean. This is confirmed by calculating the probability of splicing as a function of time, which shows that the probability function peaks more closely around the mean when the number of limiting steps increases (Fig. S2 F). To provide a quantitative measure of the variation among individual molecules, we defined the kinetic noise as the standard deviation of the splicing time, for a population of molecules, divided by the mean splicing time. The splicing noise is inversely proportional to the root square of the number of steps, and thus decreases with additional steps (see Materials and methods).

Regulation of alternative splicing

There are many ways to control alternative splicing, but one that is particularly relevant here is the so-called "kinetic control," where the time taken to transcribe the DNA between two competing splice sites creates a delay that favors splicing at the upstream site (de la Mata et al., 2003). By inducing polymerase pausing or slowing down its elongation rate, it is then possible to further increase splicing at the first site. To analyze how the kinetic noise impacts this regulation, we considered a hypothetical case of alternative splicing, where a weak acceptor site is in competition with a strong downstream site (Fig. 4 A). We assumed that splicing occurs at the upstream site only when splicing is completed before the polymerase reaches the downstream site, and that splicing must switch from 25% at the upstream acceptor site to 75%. Using the kinetic parameters of the MINX intron, a

to the bleach. (C) Scheme describing the kinetic models. (top) A single step is rate limiting (k_{spl}; red). (bottom) Multiple steps are rate-limiting, all with the same rate (k_{spl}; red). (D) Determination of the number of limiting steps. Plot of the AIC values for the optimal choice of the rate of the limiting steps (α) as a function of the number of limiting steps. L is the linear model that was included as a null hypothesis. (E) Optimal fit of the model to the experimental curve. The FRAP curve of WT_MS2in cells was fitted with either a single-step model (red), a three-step model (green), or a nine-step model (blue).

Published May 30, 2011

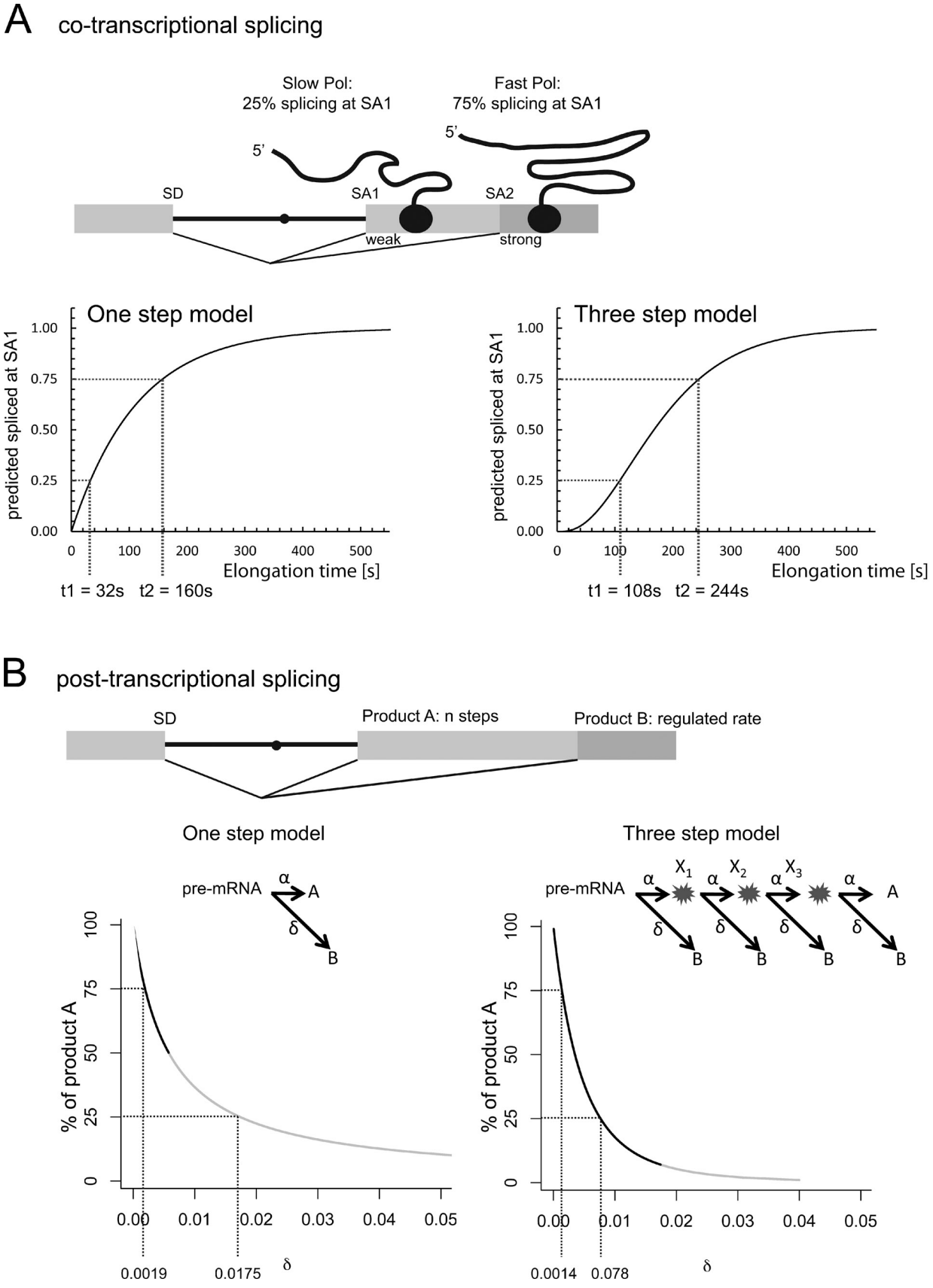

Figure 4. **A three-step model improves the regulation of alternative splicing.** (A) A three-step model improves the regulation of cotranscriptional splicing events. (A, top) Scheme of a cotranscriptional alternative splicing event regulated by the elongation rate of RNA polymerase II. Splicing occurs at splice acceptor SA1 only when the splicing reaction is completed before splice acceptor SA2 is synthesized. When both SA2 and SA1 are present on the

Published May 30, 2011

one-step model requires a fivefold reduction of the rate of polymerase elongation, from 32 to 160 s (Fig. 4 A). In contrast, because each molecule is more synchronous in a three-step model, a twofold reduction would be sufficient to achieve the same effect.

Splicing regulation by elongation rates is only one of many ways to regulate splicing. Indeed, many splicing events are regulated by cis-acting splicing factors (Wahl et al., 2009), and alternative splicing regulation can often be reproduced in vitro in the absence of transcription. It was therefore of interest to compare the impact of the one-step and three-step models on the regulation of splicing for cases where splicing regulation is disconnected from transcription. For this purpose, we considered a case of alternative splicing that does not involve transcription (Fig. 4 B). In this model, one splicing event occurs at a constitutive rate and leads to product A, whereas a competing event B occurs at a rate δ; this rate can be regulated, for instance by the expression of a key splicing factor (Fig. 4 B). We calculated the variation of rate δ required to shift the production of A from 75% to 25%, for the cases where A is made with a one-step or a three-step model. Again, the three-step model allowed for a more efficient splicing regulation, as δ has to increase only fivefold, compared to nearly tenfold for a single-step model. The three-step model also allows for a greater range of splicing regulation. Indeed, having the regulated rate δ many times greater than the constitutive rate is not plausible in a kinetic competition model, and for $\delta \leq \alpha$, the production of A can vary from 6% to 100% for a three-step model, but remains <50% for a one-step model.

Having several limiting steps instead of one thus makes the splicing reaction less noisy and allows for a greater and more efficient control of alternative splicing, where a small change in the rate of a competing process can produce a large change in splicing patterns. This is similar to the "molecular memory" model previously proposed in a theoretical study of transcription and RNA degradation (Pedraza and Paulsson, 2008). Many RNA maturation steps require the step-wise assembly of multicomponent processing machineries. Given that nascent pre-mRNAs are present as few molecules per cell, a constant time for each of the RNA processing reactions may reduce kinetic noise, and facilitate both integration of these reactions into an efficient production chain and their cross-talk for regulatory events. Many other biological pathways deal with few molecules at a time, and our model provides a general framework to explain how parallel steps in such pathways can be kinetically controlled, allowing their cross-regulation and their integration in a determined sequence of events.

Materials and methods

Cells and plasmids

The constructs were based on the MINX intron (Zillmann et al., 1988), which was fused to the LacZ gene and the 3′ end processing signals of the bovine growth hormone gene (*bGH*). The MS2 ×4 repeats were inserted within the MINX intron (WT_MS2in and SAmut_MS2in) or within exon 2 (WT_MS2ex2). These plasmids had the backbone of pcDNA3 (Invitrogen), but had the cytomegalovirus promoter replaced by the HIV-LTR. Plasmids expressing Tat have been described previously (Molle et al., 2009). GFP-MS2nsl_nes was created with the Gateway system (Invitrogen).

U2OS cells were cultivated at 37°C in DME (Invitrogen) containing 10% FCS. Stable clones were generated as described previously (Boireau et al., 2007). The copy number of the WT_MS2in reporter gene was 18–20 copies, as evaluated by quantitative PCR against a cell line having a single copy of a LacZ transgene. The number of unspliced MINX pre-mRNA at the transcription site (25 ± 10 copies) was measured by single-molecule FISH. These relatively large numbers ensure that stochastic variations in transcription initiation are averaged and that the system is at steady state. In all experiments, a Tat-expressing plasmid was transfected. Transient transfections were performed using Effectene (QIAGEN) and cells were analyzed 16–24 h later. For live cell experiments, cells were seeded on glass-bottom dishes and analyzed at 37°C in a phenol red and riboflavin-free media, supplemented with serum and 20 mM Hepes, pH 7.4. To inhibit splicing, cells were treated with 30 ng/ml SSA (Kaida et al., 2007).

In situ hybridization and immunofluorescence

Antibody concentrations were: rabbit anti-RPB1 (Euromedex) 1:3,000; anti–rabbit FITC 1:1,000 (Jackson ImmunoResearch Laboratories), and anti–rabbit Alexa Fluor 647 1:500 (Invitrogen). The anti-SF3b was a gift of R. Lührmann (Max-Planck Institute, Göttingen, Germany).

Sequences of the FISH probes were: LacZ-3, 5′-AXGAGTTTACGTGCAXGGATCTGCAACATGXCCCAGGTGACGATGTAXA-3′; LacZ-4, 5′-AXCGGCAATAATGCCXTTCCATTGTTCAGAAGGCAXCAGTCGGCTTGXA-3′; LacZ-5, 5′-AXTCAGTTCGAGGTGCXGTTTCTGGTCTXCACCCACCGGTACCAGAXA-3′; LacZ-6, 5′-AXGGGATCGATCXCGCCATACAGCGCGTXGAAACGCTGGGCAATATXA-3′; LacZ-7, 5′-AXGTTGGGTAACGCCAGGGXTTTCCCAGTCACGACGXTGTAAAACGAX-3′; LacZ-8, 5′-AXTCTTCGCTATXACGCCAGCTGGCGAAAGGGGGAXGTGCTGCAAGXA-3′; hU1, 5′-AXGGAAAACCACCXTCGTGATCATGGTATCXCCCCTGCCAGGTAAGTXA-3′; hU2, 5′-AXCCATTTAATATATTGTCCTCGGAXAGAGGACGTATCAGATATTAAACXA-3′; hU4, 5′-XXGGGAAAAGTTTXCAATTAGCAATAAXCGCGCCTCGGXX-3′; hU5, 5′-AXCTTTGATGAAAGGCGAAAGATTTAXACGATCTGAAGAGAAACCAGAGXA-3′; hU6, 5′-AXGTCATCCTTGCGCAGGGGCCAXGCTAATCTTCTCTGTATCGTTCCAAXA-3′; MS2, 5′-AXGTCGACCTGCAGACAXGGGTGATCCTCAXGTTTTCTAGGCAATXA-3′; and MINX-spliced, 5′-AXCCXCAACCGCGAGCGXTCGGAGGCCGAXA-3′.

X designates aminoallyl-modified thymidines, which were coupled with Alexa Fluor 488 or 647 (Invitrogen) or Cy3 or Cy5 (GE Healthcare) according to the manufacturer's instructions.

In situ hybridization and immunofluorescence were performed as described previously (Schmidt et al., 2006; Boireau et al., 2007). Probes for snRNAs and LacZ were hybridized in 40% formamide (Merck), whereas oligo-dT and MS2 were hybridized in 20% formamide. FISH of the spliced probe was performed in 20% formamide and washed three times in 20% formamide/1× SSC at 65°C, followed by three washes in 1× SSC at 65°C.

Microscopy

Fluorescent images of fixed cells were captured either on a MetaLsm510 confocal microscope (100×, NA 1.4; Carl Zeiss) or on a DMRA wide-field microscope (100×, NA 1.4; Leica), equipped with a CoolSNAP HQ camera (Roper Scientific) and controlled by MetaMorph (Universal Imaging). Images were prepared with Photoshop (Adobe).

pre-mRNA, splicing always occurs at SA2. When the polymerase has a fast elongation rate, 25% of the splicing events take place at site SA1, whereas when the polymerase elongates slowly, 75% of the pre-mRNAs are spliced at SA1. (A, bottom) Calculated curves of the appearance of the spliced product as a function of time after transcription (seconds), for models having a single (left) or three limiting steps (right), and optimized to fit the MINX kinetics. The vertical lines point to the elongation time required to obtain 25% and 75% of splicing at SA1. (B) A three-step model improves the regulation of posttranscriptional splicing events. (B, top) Schematics of a posttranscriptional alternative splicing event. Product A is produced constitutively with n rate-limiting steps, whereas product B is produced with a single rate-limiting step that is regulated. (B, bottom) Calculated curves of the appearance of the spliced product A as a function of the rate of the competing splicing event B, for models having a single (left) or three limiting steps (right), and optimized to fit the MINX kinetics. The vertical lines point to the rates δ required to obtain 25% and 75% of product A.

Published May 30, 2011

FRAP was performed on a microscope (TE200; Nikon) with a 100×, NA 1.45 objective lens as described previously (Boireau et al., 2007). Bleaching time was 1.2 s and recovery was monitored for 11 min with three post-bleach durations: 3D stacks were acquired every 3 s over the first 540 s, then every 6 s over 60 s, and finally every 12 s for 60 s. Recoveries at the transcription sites were then tracked in a small parallelepiped (2 × 2 × 1.5 µm) placed at the most intense area of the transcription site.

Modeling and optimization

The recovery curves of 34 cells were averaged and were normalized between 1 and 0 using the prebleached values and the 9-s time point, respectively. The process of transcription and splicing can be modeled with a set of differential equations, but to obtain optimal values for the splicing rate for a given number of steps, we have used a simplified model where transcription from the MS2 site to the 3′ splice site is simply described as a time offset (Fig. S2). We, however, verified that the optimal parameters obtained with the simplified model allowed to correctly simulate the experimental data with the full model (Fig. S2). The purpose of the fitting procedure is to determine the number of limiting step of the splicing reaction (n), and the data were thus fitted with a series of models corresponding to one and up to nine limiting steps of identical rates.

The simplified model is derived as follows. First, n species $X_1 \ldots X_n$ are introduced, such that X_1 accumulates at the transcription rate k_{tx} and is depleted at the splicing rate α, and all other X_i accumulate and deplete at rate α:

$$dX_1(t)/dt = k_{tx} - \alpha X_1$$

$$dX_n(t)/dt = \alpha X_{n-1} - \alpha X_n,\ n \geq 2.$$

The solution of this model is:

$$X_1(t) = (k_{tx}/\alpha)(1 - e^{-\alpha t})$$

$$X_n(t) = (k_{tx}/\alpha)(1 - e^{-\alpha t}) - \sum_{i=1}^{n-1} \frac{\alpha^{i-1} k_{tx} t^i e^{-\alpha t}}{i!},\ n \geq 2.$$

On rearrangement:

$$X_n(t) = (k_{tx}/\alpha)(1 - \sum_{i=0}^{n-1} \frac{\alpha^i t^i e^{-\alpha t}}{i!}),\ n \geq 1.$$

The model can then be normalized to converge to 1, as t tends to ∞ by setting $k_{tx} = \alpha$. This model is illustrated in Fig. S2 C. In this way, the model and the data are aligned in scale, and we do not need to know k_{tx} in order to find the rate α that governs the kinetics of recovery. Strictly speaking, we solve for X'_n as defined below:

$$X_n(t) = (k_{tx}/\alpha)X'_n(t),\ n \geq 1$$

As each intermediate species contributes equally to the recovered signal in proportion to the species concentration, the signal generated by n intermediates, $S_n(t)$, is given by the sum of $X_1 \ldots X_n$:

$$S_n(t) = 1 - \sum_{i=0}^{n-1} \frac{(n-i)\alpha^i t^i e^{-\alpha t}}{n i!} \quad (1)$$

For example, the signal for a process with three intermediates is $S_3(t) = X_1/3 + X_2/3 + X_3/3$.

Eq. 1 is optimized to the FRAP data by introducing an offset (β) to account for the delay in the start of recovery caused by the time required to transcribe the 3′ end of the intron. The optimization parameter k allows for uncertainty in the amplitude of the recovery. We optimize S_3 as follows:

$$S_3(t) = 1 - k[e^{-\alpha(t-\beta)} + 2\alpha(t-\beta)e^{-\alpha(t-\beta)}/3 + \alpha^2(t-\beta)^2 e^{-\alpha(t-\beta)}/6].$$

An alternative technique for optimizing the models is to allow the model to converge to a value γ other than 1 by fitting:

$$S_3(t) = \gamma - e^{-\alpha(t-\beta)} - 2\alpha(t-\beta)e^{-\alpha(t-\beta)}/3 - \alpha^2(t-\beta)^2 e^{-\alpha(t-\beta)}/6.$$

The two alternative approaches give essentially the same result, and this is whether they are fitted to FRAP data normalized between 0 and 1, or between post-bleach and 1.

To test for the possibility that the exponential models are indistinguishable because they give equally poor fits to the data, a linear model ($S_L = mt + c$) is included in the analysis as a null hypothesis. We found that this case could arise if the data were sufficiently noisy by considering synthetic data generated by stochastic simulation. Analysis indicated that our modeling and optimization techniques recover the correct number of steps (i.e., the number used to generate the data) under conditions similar to those of the experimental data, and that under noisier conditions the linear model emerges as the best model. Thus, we have confidence that the number of steps identified by our method is a realistic reflection of the underlying process.

Optimization was performed using the nonlinear least squares (nls) package in R, and models were compared using the associated AIC method, which can compute P-values for each model (Burnham and Anderson, 2002). We also evaluated the goodness of the fits by analyzing the residuals and computing the probability that their distribution corresponds to a normal function with a Shapiro-Wilk test. Finally, we compared the precision of the fit to the experimental error using the reduced χ^2. All these indicators are given in Table S1 and consistently give a model with three limiting steps as the most probable model for splicing. Obtained time constants were then used to calculate the appearance of the spliced product over time.

Kinetic noise

Keeping the mean reaction time T constant, the Erlang distribution for an integer values of n (the number of steps) gives $n/lambda = T$ (Weber et al., 2005). Then the variance is $n/lambda^2$, or T^2/n, and the kinetic noise is SD/mean = $1/\sqrt{n}$.

Co-transcriptional splicing assay

Total RNA was isolated using TRIZOL (Invitrogen) followed by phenol-chloroform extraction and isopropanol precipitation according to the manufacturer's instructions. For reverse transcription, 1–5 µg of total RNA was treated with RNase-free DNaseI (Promega) and transcribed using the Im-PromII Reverse Transcription kit (Promega) either with a poly-dT primer or a 3′ uncleaved primer, hybridizing to the sequence beyond the bGH polyA cleavage site. The sequence of the uncleaved primer for the MINX reporter was: 5′-GAATGACACCTACTCAGA-3′. To detect unspliced and spliced fragments at the same time, PCR primers binding to exon 1, 5′-TCTAGCAGTGGCGCCGAAC-3′, and exon 2, 5′-GCAACATGTCCCA-GGTGACGATGTATTTT-3′, were used.

For the endogenous genes of Table I, RT was primed either with oligodT or a gene-specific primer hybridizing after the 3′ end cleavage site, and splicing was assessed using a competitive PCR assay with three primers, two that hybridize in the exons flanking the intron, and one that binds the intron, as depicted in Fig. 1 A. The size of the "spliced" and "unspliced" PCR products was similar to avoid amplification bias. PCR reactions were stopped at different cycles, run on gels, and quantified after EtBr staining. For alternative splicing events, ratios of isoforms were also quantified by PCR using oligo-dT as an RT primer, and the percentage of cotranscriptional splicing was normalized to the percentage of intron inclusion in polyA mRNAs.

Online supplemental material

Fig. S1 shows Northern blot analysis of cells stably expressing WT_MS2in and SAmut_MS2in reporter RNAs. Fig. S2 shows a model of transcription and splicing, and fits to the experimental data. Table S1 shows optimal model parameters that fit the experiment intronic FRAP curve. Online supplemental material is available at http://www.jcb.org/cgi/content/full/jcb.201009012/DC1.

We thank R. Bordonné and S. Rader for critical reading of the manuscript.

This work was supported by Sidaction, Agence Nationale de Recherche sur le Sida, the EC RiboSys (grant LSHG-CT-2005-518280), and EURASNET. S. Aitken was supported by the Centre for Systems Biology at Edinburgh, a Centre for Integrative Systems Biology funded by the Biotechnology and Biological Sciences Research Council and the Engineering and Physical Sciences Research Council (BB/D019621/1), and by a Wellcome Trust Value In People award. U. Schmidt was supported by EMBO (171-2006) and the Deutsche Forschungsgemeinschaft (Schm 2659/1-1).

Submitted: 7 September 2010
Accepted: 25 April 2011

Published May 30, 2011

References

Audibert, A., D. Weil, and F. Dautry. 2002. In vivo kinetics of mRNA splicing and transport in mammalian cells. *Mol. Cell. Biol.* 22:6706–6718. doi:10.1128/MCB.22.19.6706-6718.2002

Baurén, G., and L. Wieslander. 1994. Splicing of Balbiani ring 1 gene pre-mRNA occurs simultaneously with transcription. *Cell*. 76:183–192. doi:10.1016/0092-8674(94)90182-1

Beyer, A.L., and Y.N. Osheim. 1988. Splice site selection, rate of splicing, and alternative splicing on nascent transcripts. *Genes Dev.* 2:754–765. doi:10.1101/gad.2.6.754

Blencowe, B.J. 2006. Alternative splicing: new insights from global analyses. *Cell*. 126:37–47. doi:10.1016/j.cell.2006.06.023

Boireau, S., P. Maiuri, E. Basyuk, M. de la Mata, A. Knezevich, B. Pradet-Balade, V. Bäcker, A. Kornblihtt, A. Marcello, and E. Bertrand. 2007. The transcriptional cycle of HIV-1 in real-time and live cells. *J. Cell Biol.* 179:291–304. doi:10.1083/jcb.200706018

Burnham, K.P., and D.R. Anderson. 2002. Model selection and multimodel inference. Springer. 488 pp.

Carrillo Oesterreich, F., S. Preibisch, and K.M. Neugebauer. 2010. Global analysis of nascent RNA reveals transcriptional pausing in terminal exons. *Mol. Cell*. 40:571–581. doi:10.1016/j.molcel.2010.11.004

Darzacq, X., Y. Shav-Tal, V. de Turris, Y. Brody, S.M. Shenoy, R.D. Phair, and R.H. Singer. 2007. In vivo dynamics of RNA polymerase II transcription. *Nat. Struct. Mol. Biol.* 14:796–806. doi:10.1038/nsmb1280

de la Mata, M., C.R. Alonso, S. Kadener, J.P. Fededa, M. Blaustein, F. Pelisch, P. Cramer, D. Bentley, and A.R. Kornblihtt. 2003. A slow RNA polymerase II affects alternative splicing in vivo. *Mol. Cell*. 12:525–532. doi:10.1016/j.molcel.2003.08.001

de la Mata, M., C. Lafaille, and A.R. Kornblihtt. 2010. First come, first served revisited: factors affecting the same alternative splicing event have different effects on the relative rates of intron removal. *RNA*. 16:904–912. doi:10.1261/rna.1993510

Fusco, D., N. Accornero, B. Lavoie, S.M. Shenoy, J.M. Blanchard, R.H. Singer, and E. Bertrand. 2003. Single mRNA molecules demonstrate probabilistic movement in living mammalian cells. *Curr. Biol.* 13:161–167. doi:10.1016/S0960-9822(02)01436-7

Huranová, M., I. Ivani, A. Benda, I. Poser, Y. Brody, M. Hof, Y. Shav-Tal, K.M. Neugebauer, and D. Stanek. 2010. The differential interaction of snRNPs with pre-mRNA reveals splicing kinetics in living cells. *J. Cell Biol.* 191:75–86. doi:10.1083/jcb.201004030

Kaida, D., H. Motoyoshi, E. Tashiro, T. Nojima, M. Hagiwara, K. Ishigami, H. Watanabe, T. Kitahara, T. Yoshida, H. Nakajima, et al. 2007. Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. *Nat. Chem. Biol.* 3:576–583. doi:10.1038/nchembio.2007.18

Kim, E., L. Du, D.B. Bregman, and S.L. Warren. 1997. Splicing factors associate with hyperphosphorylated RNA polymerase II in the absence of pre-mRNA. *J. Cell Biol.* 136:19–28. doi:10.1083/jcb.136.1.19

Kiseleva, E., T. Wurtz, N. Visa, and B. Daneholt. 1994. Assembly and disassembly of spliceosomes along a specific pre-messenger RNP fiber. *EMBO J.* 13:6052–6061.

McCracken, S., N. Fong, K. Yankulov, S. Ballantyne, G. Pan, J. Greenblatt, S.D. Patterson, M. Wickens, and D.L. Bentley. 1997. The C-terminal domain of RNA polymerase II couples mRNA processing to transcription. *Nature*. 385:357–361. doi:10.1038/385357a0

Molle, D., C. Segura-Morales, G. Camus, C. Berlioz-Torrent, J. Kjems, E. Basyuk, and E. Bertrand. 2009. Endosomal trafficking of HIV-1 gag and genomic RNAs regulates viral egress. *J. Biol. Chem.* 284:19727–19743. doi:10.1074/jbc.M109.019844

Muñoz, M.J., M.S. Pérez Santangelo, M.P. Paronetto, M. de la Mata, F. Pelisch, S. Boireau, K. Glover-Cutter, C. Ben-Dov, M. Blaustein, J.J. Lozano, et al. 2009. DNA damage regulates alternative splicing through inhibition of RNA polymerase II elongation. *Cell*. 137:708–720. doi:10.1016/j.cell.2009.03.010

Pandit, S., D. Wang, and X.D. Fu. 2008. Functional integration of transcriptional and RNA processing machineries. *Curr. Opin. Cell Biol.* 20:260–265. doi:10.1016/j.ceb.2008.03.001

Pandya-Jones, A., and D.L. Black. 2009. Co-transcriptional splicing of constitutive and alternative exons. *RNA*. 15:1896–1908. doi:10.1261/rna.1714509

Pedraza, J.M., and J. Paulsson. 2008. Effects of molecular memory and bursting on fluctuations in gene expression. *Science*. 319:339–343. doi:10.1126/science.1144331

Roberts, G.C., C. Gooding, H.Y. Mak, N.J. Proudfoot, and C.W. Smith. 1998. Co-transcriptional commitment to alternative splice site selection. *Nucleic Acids Res.* 26:5568–5572. doi:10.1093/nar/26.24.5568

Roybal, G.A., and M.S. Jurica. 2010. Spliceostatin A inhibits spliceosome assembly subsequent to prespliceosome formation. *Nucleic Acids Res.* 38:6664–6672. doi:10.1093/nar/gkq494

Schmidt, U., K. Richter, A.B. Berger, and P. Lichter. 2006. In vivo BiFC analysis of Y14 and NXF1 mRNA export complexes: preferential localization within and around SC35 domains. *J. Cell Biol.* 172:373–381. doi:10.1083/jcb.200503061

Singh, J., and R.A. Padgett. 2009. Rates of in situ transcription and splicing in large human genes. *Nat. Struct. Mol. Biol.* 16:1128–1133. doi:10.1038/nsmb.1666

Sultan, M., M.H. Schulz, H. Richard, A. Magen, A. Klingenhoff, M. Scherf, M. Seifert, T. Borodina, A. Soldatov, D. Parkhomchuk, et al. 2008. A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science*. 321:956–960. doi:10.1126/science.1160342

Tardiff, D.F., S.A. Lacadie, and M. Rosbash. 2006. A genome-wide analysis indicates that yeast pre-mRNA splicing is predominantly posttranscriptional. *Mol. Cell*. 24:917–929. doi:10.1016/j.molcel.2006.12.002

Wada, Y., Y. Ohta, M. Xu, S. Tsutsumi, T. Minami, K. Inoue, D. Komura, J. Kitakami, N. Oshida, A. Papantonis, et al. 2009. A wave of nascent transcription on activated human genes. *Proc. Natl. Acad. Sci. USA*. 106:18357–18361. doi:10.1073/pnas.0902573106

Wahl, M.C., C.L. Will, and R. Lührmann. 2009. The spliceosome: design principles of a dynamic RNP machine. *Cell*. 136:701–718. doi:10.1016/j.cell.2009.02.009

Wang, E.T., R. Sandberg, S. Luo, I. Khrebtukova, L. Zhang, C. Mayr, S.F. Kingsmore, G.P. Schroth, and C.B. Burge. 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature*. 456:470–476. doi:10.1038/nature07509

Weber, A., J. Liu, I. Collins, and D. Levens. 2005. TFIIH operates through an expanded proximal promoter to fine-tune c-myc expression. *Mol. Cell. Biol.* 25:147–161. doi:10.1128/MCB.25.1.147-161.2005

Wetterberg, I., J. Zhao, S. Masich, L. Wieslander, and U. Skoglund. 2001. In situ transcription and splicing in the Balbiani ring 3 gene. *EMBO J.* 20:2564–2574. doi:10.1093/emboj/20.10.2564

Zhang, G., K.L. Taneja, R.H. Singer, and M.R. Green. 1994. Localization of pre-mRNA splicing in mammalian nuclei. *Nature*. 372:809–812.

Zillmann, M., M.L. Zapp, and S.M. Berget. 1988. Gel electrophoretic isolation of splicing complexes containing U1 small nuclear ribonucleoprotein particles. *Mol. Cell. Biol.* 8:814–821.